



## **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/884,752  
Source: OLP  
Date Processed by STIC: 2/5/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

## Raw Sequence Listing Error Summary

### ERROR DETECTED

### SUGGESTED CORRECTION

SERIAL NUMBER: 09/88/752

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos     The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
  
- 2      Invalid Line Length     The rules require that a line not exceed 72 characters in length. This includes white spaces.
  
- 3      Misaligned Amino  
    Numbering     The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
  
- 4      Non-ASCII     The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
  
- 5      Variable Length     Sequence(s)          contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
  
- 6      PatentIn 2.0  
    "bug"     A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)         . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
  
- 7      Skipped Sequences  
    (OLD RULES)     Sequence(s)          missing. If intentional, please insert the following lines for each skipped sequence:  
                               (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                               (i)     SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                               (xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                               This sequence is intentionally skipped  
  
                               Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
  
- 8      Skipped Sequences  
    (NEW RULES)     Sequence(s)          missing. If intentional, please insert the following lines for each skipped sequence.  
                               <210> sequence id number  
                               <400> sequence id number  
                               000
  
- 9      Use of n's or Xaa's  
    (NEW RULES)     Use of n's and/or Xaa's have been detected in the Sequence Listing.  
                               Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
                               In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
  
- 10     Invalid <213>  
    Response     Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
  
- 11     Use of <220>     Sequence(s)          missing the <220> "Feature" and associated numeric identifiers and responses.  
                               Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                               (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
  
- 12     PatentIn 2.0  
    "bug"     Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

OIPE

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:39

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

Does Not Comply  
Corrected Diskette Needed

## SEQUENCE LISTING

## C--&gt; 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Harold Kleanthous et al.

## C--&gt; 7 (ii) TITLE OF INVENTION: Identification of Polynucleotides

8 Encoding Novel Helicobacter Polypeptides in the

Helicobacter

9 Genome

11 (iii) NUMBER OF SEQUENCES: 370

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Clark &amp; Elbing LLP

15 (B) STREET: 176 Federal Street

16 (C) CITY: Boston

17 (D) STATE: MA

18 (E) COUNTRY: USA

19 (F) ZIP: 02110

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Diskette

23 (B) COMPUTER: IBM Compatible

24 (C) OPERATING SYSTEM: DOS

25 (D) SOFTWARE: FastSEQ for Windows Version 2.0

27 (vi) CURRENT APPLICATION DATA:

C--&gt; 28 (A) APPLICATION NUMBER: US/09/881,752

C--&gt; 29 (B) FILING DATE: 18-Jun-2001

30 (C) CLASSIFICATION:

32 (viii) ATTORNEY/AGENT INFORMATION:

33 (A) NAME: Clark, Paul T.

34 (B) REGISTRATION NUMBER: 30,162

35 (C) REFERENCE/DOCKET NUMBER: 06132/041001

37 (ix) TELECOMMUNICATION INFORMATION:

38 (A) TELEPHONE: 617-428-0200

39 (B) TELEFAX: 617-428-7045

40 (C) TELEX:

## ERRORED SEQUENCES

1361 (2) INFORMATION FOR SEQ ID NO: 21:

1363 (i) SEQUENCE CHARACTERISTICS:

1364 (A) LENGTH: 2169 base pairs

1365 (B) TYPE: nucleic acid

1366 (C) STRANDEDNESS: single

1367 (D) TOPOLOGY: linear

1369 (ii) MOLECULE TYPE: Genomic DNA

1370 (ix) FEATURE:

1372 (A) NAME/KEY: Coding Sequence

1373 (B) LOCATION: 60...2039

1374 (D) OTHER INFORMATION:

1376 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

P.4

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:39

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

1378	CCCTATCATA GGGCGTGGCA TGAAGAAAAA AGCAAAAGTC TTTTGGTATT GTTTTAATC	59
1381	ATG ATT TAT TGG TTG TAT TTG GCG GTC TTT TTT TTG TTG AGC GCA TTA	107
1382	Met Ile Tyr Trp Leu Tyr Leu Ala Val Phe Phe Leu Leu Ser Ala Leu	
1383	1 5 10 15	
1385	GAC GCT AAA GAA ATC GCT ATG CAA CGA TTT GAC AAA CAA AAC CAT AAG	155
1386	Asp Ala Lys Glu Ile Ala Met Gln Arg Phe Asp Lys Gln Asn His Lys	
1387	20 25 30	
1389	ATT TTT GAA ATC CTT GCG GAT AAA GTG AGC GCT AAA GAC AAT GTG ATA	203
1390	Ile Phe Glu Ile Leu Ala Asp Lys Val Ser Ala Lys Asp Asn Val Ile	
1391	35 40 45	
1393	ACC GCA TCA GGG AAT GCG ATC TTA TTG AAT TAT GAT GTG TAT ATT CTA	251
1394	Thr Ala Ser Gly Asn Ala Ile Leu Leu Asn Tyr Asp Val Tyr Ile Leu	
1395	50 55 60	
1397	GCG GAC AAG GTG CGT TAT GAC ACT AAA ACC AAA GAA GCG TTA TTA GAG	299
1398	Ala Asp Lys Val Arg Tyr Asp Thr Lys Thr Lys Glu Ala Leu Leu Glu	
1399	65 70 75 80	
1401	GGG AAT ATC AAG GTT TAT AGG GGC GAG GGT TTG CTC GTT AAA ACC GAT	347
1402	Gly Asn Ile Lys Val Tyr Arg Gly Glu Gly Leu Leu Val Lys Thr Asp	
1403	85 90 95	
1405	TAC GTG AAA TTG AGT TTG AAT GAA AAA TAT GAA ATC ATT TTC CCC TTT	395
1406	Tyr Val Lys Leu Ser Leu Asn Glu Lys Tyr Glu Ile Ile Phe Pro Phe	
1407	100 105 110	
1409	TAT GTC CAA GAC AGC GTG AGC GGG ATT TGG GTG AGC GCG GAT ATT GCC	443
1410	Tyr Val Gln Asp Ser Val Ser Gly Ile Trp Val Ser Ala Asp Ile Ala	
1411	115 120 125	
1413	AGC GGA AAG GAT CAA AAA TAT AAG GTT AAA AAC ATG AGC ACT TCA GGG	491
1414	Ser Gly Lys Asp Gln Lys Tyr Lys Val Lys Asn Met Ser Thr Ser Gly	
1415	130 135 140	
1417	TGC AGC ATT GAT AAC CCC ATT TGG CAT GTC AAT GCG ACT TCA GGC TCA	539
1418	Cys Ser Ile Asp Asn Pro Ile Trp His Val Asn Ala Thr Ser Gly Ser	
1419	145 150 155 160	
1421	TTC AAC ATG CAA AAA TCG CAT TTG TCT ATG TGG AAT CCT AAG ATC TAT	587
1422	Phe Asn Met Gln Lys Ser His Leu Ser Met Trp Asn Pro Lys Ile Tyr	
1423	165 170 175	
1425	GTC GGT GAT ATT CCT GTA TTG TAT TTG CCC TAT ATT TTC ATG TCC ACG	635
1426	Val Gly Asp Ile Pro Val Leu Tyr Leu Pro Tyr Ile Phe Met Ser Thr	
1427	180 185 190	
1429	AGC AAT AAA AGA ACT ACT GGG TTT TTA TAC CCT GAG TTT GGC ACT TCC	683
1430	Ser Asn Lys Arg Thr Thr Gly Phe Leu Tyr Pro Glu Phe Gly Thr Ser	
1431	195 200 205	
1433	AAC TTA GAC GGC TTT ATT TAT TTG CAA CCC TTT TAT TTA GCC CCC AAA	731
1434	Asn Leu Asp Gly Phe Ile Tyr Leu Gln Pro Phe Tyr Leu Ala Pro Lys	
1435	210 215 220	
1437	AAC TCA TGG GAT ATG ACC TTT ACC CCA CAA ATC CGC TAT AAA AGG GGT	779
1438	Asn Ser Trp Asp Met Thr Phe Thr Pro Gln Ile Arg Tyr Lys Arg Gly	
1439	225 230 235 240	
1441	TTT GGC TTG AAT TTT GAA GCG CGC TAC ATT AAC TCT AAA AAC GAC AGG	827
1442	Phe Gly Leu Asn Phe Glu Ala Arg Tyr Ile Asn Ser Lys Asn Asp Arg	
1443	245 250 255	

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:39

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

1445	TTT	TTA	TTC	AAC	GCG	CGC	TAT	TTT	AGG	AAT	TAC	ACC	CAA	TAT	GTC	AAA	875
1446	Phe	Leu	Phe	Asn	Ala	Arg	Tyr	Phe	Arg	Asn	Tyr	Thr	Gln	Tyr	Val	Lys	
1447				260					265					270			
1449	CGC	TAC	GAT	TTG	AGG	AAT	CAA	AAT	ATC	TAC	GGG	TTT	GAA	TTT	TTA	AGC	923
1450	Arg	Tyr	Asp	Leu	Arg	Asn	Gln	Asn	Ile	Tyr	Gly	Phe	Glu	Phe	Leu	Ser	
1451				275					280					285			
1453	TCT	AGC	AGG	GAC	ACT	TTA	CAA	AAA	TAC	TTC	CAC	CTT	AAG	TCT	AAT	ATT	971
1454	Ser	Ser	Arg	Asp	Thr	Leu	Gln	Lys	Tyr	Phe	His	Leu	Lys	Ser	Asn	Ile	
1455				290					295				300				
1457	GAC	AAC	GGG	CAT	TAC	ATT	GAC	TTT	TTA	TAC	ATG	AAC	GAT	TTG	GAC	TAT	1019
1458	Asp	Asn	Gly	His	Tyr	Ile	Asp	Phe	Leu	Tyr	Met	Asn	Asp	Leu	Asp	Tyr	
1459				305					310				315			320	
1461	GTG	CGT	TTT	GAA	AAG	GTT	AAT	AAG	CGT	ATC	ACA	GAC	GCC	ACG	CAC	ATG	1067
1462	Val	Arg	Phe	Glu	Lys	Val	Asn	Lys	Arg	Ile	Thr	Asp	Ala	Thr	His	Met	
1463				325					330					335			
1465	TCT	AGG	GCG	AAT	TAC	TAT	TTG	CAA	ACA	GAA	AAC	AAT	TAT	TAC	GGC	TTG	1115
1466	Ser	Arg	Ala	Asn	Tyr	Tyr	Leu	Gln	Thr	Glu	Asn	Asn	Tyr	Tyr	Gly	Leu	
1467				340					345					350			
1469	AAT	ATC	AAG	TAT	TTT	TTA	AAC	CTG	AAT	AAA	ATC	AAC	AAT	AAC	CGC	ACT	1163
1470	Asn	Ile	Lys	Tyr	Phe	Leu	Asn	Leu	Asn	Lys	Ile	Asn	Asn	Asn	Arg	Thr	
1471				355					360					365			
1473	TTC	CAA	TCT	GTC	CCT	AAT	TTG	CAA	TAC	CAT	AAA	TAT	TTA	AAT	TCT	TTG	1211
1474	Phe	Gln	Ser	Val	Pro	Asn	Leu	Gln	Tyr	His	Lys	Tyr	Leu	Asn	Ser	Leu	
1475				370					375				380				
1477	TAT	TTT	AGA	AAT	TTG	TTG	TAT	TCG	GTG	GAT	TAT	CAG	TTT	AGA	AAC	ACC	1259
1478	Tyr	Phe	Arg	Asn	Leu	Leu	Tyr	Ser	Val	Asp	Tyr	Gln	Phe	Arg	Asn	Thr	
1479				385					390				395			400	
1481	GCA	AGA	GAG	ATT	GGT	TAT	GGC	TAT	GTG	CAA	AAC	GCT	TTG	AAT	GTG	CCG	1307
1482	Ala	Arg	Glu	Ile	Gly	Tyr	Gly	Tyr	Val	Gln	Asn	Ala	Leu	Asn	Val	Pro	
1483				405					410					415			
1485	GTG	GGC	TTG	CAA	TTT	TCT	TTG	TTT	AAA	AAG	TAT	TTG	TCT	TTA	GGG	CTT	1355
1486	Val	Gly	Leu	Gln	Phe	Ser	Leu	Phe	Lys	Lys	Tyr	Leu	Ser	Leu	Gly	Leu	
1487				420					425					430			
1489	TGG	AAT	GAT	CTC	CAA	CTA	TCT	AAT	GTG	GCT	TTA	ATG	CAA	TCT	AAA	AAT	1403
1490	Trp	Asn	Asp	Leu	Gln	Leu	Ser	Asn	Val	Ala	Leu	Met	Gln	Ser	Lys	Asn	
1491				435					440					445			
1493	TCC	TTC	GTG	CCT	ACG	ATC	CCT	AAT	GAA	TCA	AGG	GAA	TTT	GGG	AAT	TTT	1451
1494	Ser	Phe	Val	Pro	Thr	Ile	Pro	Asn	Glu	Ser	Arg	Glu	Phe	Gly	Asn	Phe	
1495				450					455					460			
1497	GTG	TCT	TCA	AAT	TTT	TCC	ATG	TAT	GTC	AAT	ACG	GAT	TTG	GCT	AGA	GAA	1499
1498	Val	Ser	Ser	Asn	Phe	Ser	Met	Tyr	Val	Asn	Thr	Asp	Leu	Ala	Arg	Glu	
1499				465					470					475		480	
1501	TAC	AAC	AAG	CTT	TTC	CAC	ACG	ATC	CAA	CTA	GAA	GCG	ATT	TTC	AAC	ATC	1547
1502	Tyr	Asn	Lys	Leu	Phe	His	Thr	Ile	Gln	Leu	Glu	Ala	Ile	Phe	Asn	Ile	
1503				485					490						495		
1505	CCT	TAT	TAC	ACC	TTT	AAA	AAC	GGC	TTA	TTT	TCT	CAA	AAC	ATG	TAT	GCT	1595
1506	Pro	Tyr	Tyr	Thr	Phe	Lys	Asn	Gly	Leu	Phe	Ser	Gln	Asn	Met	Tyr	Ala	
1507				500					505					510			
1509	TTA	AGC	GCG	CAA	GCC	TTA	AAC	AGC	TAC	ACT	TCG	CCT	TTA	TTG	AGA	GAT	1643

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/881,752

DATE: 07/05/2001  
TIME: 15:54:39

Input Set : A:\06132.041001 sequence listing.txt  
Output Set: N:\CRF3\07032001\I881752.raw

1510 Leu Ser Ala Gln Ala Leu Asn Ser Tyr Thr Ser Pro Leu Leu Arg Asp  
1511 515 520 525  
1513 TAT GAT TAT CAA GGG CGT TTG TAT GAC TCG GTG TGG AAT CCT AGC AGT 1691  
1514 Tyr Asp Tyr Gln Gly Arg Leu Tyr Asp Ser Val Trp Asn Pro Ser Ser  
1515 530 535 540  
1517 ATT TTA CCT AGC AAT GCG AGC AAC AAG ACG GTG GAT TTA ACC CTA ACG 1739  
1518 Ile Leu Pro Ser Asn Ala Ser Asn Lys Thr Val Asp Leu Thr Leu Thr  
1519 545 550 555 560  
1521 CAA TAC CTT TAT GGC TTA GGG GGG CAA GAG TTA TTG TAT TTT AAA ATA 1787  
E--> 1522 Gln Tyr Leu Tyr Gly Leu Gly Gly Gln Glu Leu Leu Tyr Phe Lys Ile *misaligned amino acid nos.*  
E--> 1524 TGG CAA CTC ATC AAT CTT GAC GAT AAA GTT TCG CCC TTT AGA ATG CCA 1835  
1525 Ser Gln Leu Ile Asn Leu Asp Asp Lys Val Ser Pro Phe Arg Met Pro  
W--> 1526 580 585 590  
E--> 1528 CTA GAG AGC AAG ATC GGG TTT TCG CCC TTA ACG GGA TTG AAC ATC TTT 1883  
1529 Leu Glu Ser Lys Ile Gly Phe Ser Pro Leu Thr Gly Leu Asn Ile Phe  
W--> 1530 595 600 605  
E--> 1532 GGG AAT GTC TTT TAT TCG TTT TAT CAA AAC CGC TTA GAA GAA ATC TCT 1931  
1533 Gly Asn Val Phe Tyr Ser Phe Tyr Gln Asn Arg Leu Glu Glu Ile Ser  
W--> 1534 610 615 620  
E--> 1536 GTG AAC GCC AAT TAC CAA CGC AAG TTT TTA AGC TTT AAC CTC TCT TAT 1979  
1537 Val Asn Ala Asn Tyr Gln Arg Lys Phe Leu Ser Phe Asn Leu Ser Tyr  
W--> 1538 625 630 635 640  
E--> 1540 TTT TTA AAA AAC AAT TTT AGC AGT GGG ATT AAT AGC ATT GTA GAA AAT 2027  
1541 Phe Leu Lys Asn Asn Phe Ser Ser Gly Ile Asn Ser Ile Val Glu Asn  
W--> 1542 645 650 655  
E--> 1544 CTG CGG ATT ATT TAAAGGCGGG TTTTAGCAAC GACTTTGGCT ATTTTTCAT GAGCGC 2085  
1545 Leu Arg Ile Ile  
W--> 1546 660  
E--> 1548 GGATGTGGGT TATGATATTA GAAACAATGT GGTTTTAAAT TGGAATGTGG GGATTTATAA 2145  
E--> 1549 AAAAATCCGT TGTTTTGGGA TTGG 2169

8311 (2) INFORMATION FOR SEQ ID NO: 111:  
8313 (i) SEQUENCE CHARACTERISTICS:  
8314 (A) LENGTH: 397 base pairs  
8315 (B) TYPE: nucleic acid  
8316 (C) STRANDEDNESS: single  
8317 (D) TOPOLOGY: linear  
8319 (ii) MOLECULE TYPE: Genomic DNA  
8320 (ix) FEATURE:  
8322 (A) NAME/KEY: Coding Sequence  
8323 (B) LOCATION: 53...352  
8324 (D) OTHER INFORMATION:  
8328 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

E--> 8330 CTAATTCTGT CTATTACACC AACAATCAAT CTCAAAACAA AGGACATGAAAG ATG AAA 58  
8331 Met Lys  
8332 1  
8334 ACA AAA CAT AAA GGA ATA AGA ATG TTT AAG CAA ATT CGT AGA ATG ATG 106  
8335 Thr Lys His Lys Gly Ile Arg Met Phe Lys Gln Ile Arg Arg Met Met  
8336 5 10 15  
8338 AGT TTG GCA ATA TTA ATG CCT AGT TTT TTA TTG GCG GCA CCA GAT TAC 154

*insert a space - Per 1.822  
of sequence  
rules,  
group all  
non-coding  
bases into  
10's*

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:40

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

8339 Ser Leu Ala Ile Leu Met Pro Ser Phe Leu Leu Ala Ala Pro Asp Tyr  
8340 20 25 30  
8342 AAA CAA AAA TTC ACT CAA ATA TTG GAT TTC ATA AGC AAT GAC TTT ATC 202  
8343 Lys Gln Lys Phe Thr Gln Ile Leu Asp Phe Ile Ser Asn Asp Phe Ile  
8344 35 40 45 50  
8346 AAG GCT ATT GGT GGT CTA ATC ATT GTT GGG ACT TGC ATT TAC GCC TAT 250  
8347 Lys Ala Ile Gly Gly Leu Ile Ile Val Gly Thr Cys Ile Tyr Ala Tyr  
8348 55 60 65  
8350 AAA AAT TGG GAC AGG CTT GGA GAA ATT GGT TGG AAA TGC GTT GGG ATT 298  
8351 Lys Asn Trp Asp Arg Leu Gly Glu Ile Gly Trp Lys Cys Val Gly Ile  
8352 70 75 80  
8354 ATC ATT ATA ACC GCT GCT ATT TCT AAT GCT AAA ACT TTA AGT CAA TGG 346  
8355 Ile Ile Ile Thr Ala Ala Ile Ser Asn Ala Lys Thr Leu Ser Gln Trp  
8356 85 90 95  
8358 TTA TTT TAGATGGCAT TGCATATTGT TTGTGTTGAA AGTATCAACA TTAGA 397  
8359 Leu Phe  
8360 100

10141 (2) INFORMATION FOR SEQ ID NO: 138:  
10143 (i) SEQUENCE CHARACTERISTICS:  
10144 (A) LENGTH: 1290 amino acids  
10145 (B) TYPE: amino acid  
10146 (C) STRANDEDNESS: single  
10147 (D) TOPOLOGY: linear  
10149 (ii) MOLECULE TYPE: protein  
10150 (v) FRAGMENT TYPE: internal  
10152 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

10154 Met Glu Ile Gln Gln Thr His Arg Lys Ile Asn Arg Pro Leu Val Ser  
10155 1 5 10 15  
10156 Leu Ala Leu Val Gly Ala Leu Val Ser Ile Thr Pro Gln Gln Ser His  
10157 20 25 30  
10158 Ala Ala Phe Phe Thr Thr Val Ile Ile Pro Ala Ile Val Gly Gly Ile  
10159 35 40 45  
10160 Ala Thr Gly Ala Ala Val Gly Thr Val Ser Gly Leu Leu Gly Trp Gly  
10161 50 55 60  
10162 Leu Lys Gln Ala Glu Glu Ala Asn Lys Thr Pro Asp Lys Pro Asp Lys  
10163 65 70 75 80  
10164 Val Trp Arg Ile Gln Ala Gly Lys Gly Phe Asn Glu Phe Pro Asn Lys  
10165 85 90 95  
10166 Glu Tyr Asp Leu Tyr Arg Ser Leu Leu Ser Ser Lys Ile Asp Gly Gly  
10167 100 105 110  
10168 Trp Asp Trp Gly Asn Ala Ala Thr His Tyr Trp Val Lys Gly Gly Gln  
10169 115 120 125  
10170 Trp Asn Lys Leu Glu Val Asp Met Lys Asp Ala Val Gly Thr Tyr Asn  
10171 130 135 140  
10172 Leu Ser Gly Leu Arg Asn Phe Thr Gly Gly Asp Leu Asp Val Asn Met  
10173 145 150 155 160  
10174 Gln Lys Ala Thr Leu Arg Leu Gly Gln Phe Asn Gly Asn Ser Phe Thr  
10175 165 170 175  
10176 Ser Tyr Lys Asp Ser Ala Asp Arg Thr Thr Arg Val Asp Phe Asn Ala





## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/881,752

DATE: 07/05/2001

TIME: 15:54:40

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

10226 Arg Ile Asn Thr Val Arg Leu Glu Thr Gly Thr Arg Ser Ile Phe Ser
10227           580           585           590
10228 Gly Gly Val Lys Phe Lys Ser Gly Glu Lys Leu Val Ile Asp Glu Phe
10229           595           600           605
10230 Tyr Tyr Ser Pro Trp Asn Tyr Phe Asp Ala Arg Asn Ile Lys Asn Val
10231           610           615           620
10232 Glu Ile Thr Arg Lys Phe Ala Ser Ser Thr Pro Glu Asn Pro Trp Gly
10233           625           630           635           640
10234 Thr Ser Lys Leu Met Phe Asn Asn Leu Thr Leu Gly Gln Asn Ala Val
10235           645           650           655
10236 Met Asp Tyr Ser Gln Phe Ser Asn Leu Thr Ile Gln Gly Asp Phe Ile
10237           660           665           670
10238 Asn Asn Gln Gly Thr Ile Asn Tyr Leu Val Arg Gly Gly Gln Val Ala
10239           675           680           685
10240 Thr Leu Asn Val Gly Asn Ala Ala Ala Met Phe Phe Ser Asn Asn Val
10241           690           695           700
10242 Asp Ser Ala Thr Gly Phe Tyr Gln Pro Leu Met Lys Ile Asn Ser Ala
10243           705           710           715           720
10244 Gln Asp Leu Ile Lys Asn Lys Glu His Val Leu Leu Lys Ala Lys Ile
10245           725           730           735
10246 Ile Gly Tyr Gly Asn Val Ser Leu Gly Thr Asn Ser Ile Ser Asn Val
10247           740           745           750
10248 Asn Leu Ile Glu Gln Phe Lys Glu Arg Leu Ala Leu Tyr Asn Asn Asn
10249           755           760           765
10250 Asn Arg Met Asp Ile Cys Val Val Arg Asn Thr Asp Asp Ile Lys Ala
10251           770           775           780
10252 Cys Gly Thr Ala Ile Gly Asn Gln Ser Met Val Asn Asn Pro Asp Asn
10253           785           790           795           800
10254 Tyr Lys Tyr Leu Ile Gly Lys Ala Trp Lys Asn Ile Gly Ile Ser Lys
10255           805           810           815
10256 Thr Ala Asn Gly Ser Lys Ile Ser Val Tyr Tyr Leu Gly Asn Ser Thr
10257           820           825           830
10258 Pro Thr Glu Lys Gly Gly Asn Thr Thr Asn Leu Pro Thr Asn Thr Thr
10259           835           840           845
10260 Ser Asn Val Arg Ser Ala Asn Asn Ala Leu Ala Gln Asn Ala Pro Phe
10261           850           855           860
10262 Ala Gln Pro Ser Ala Thr Pro Asn Leu Val Ala Ile Asn Gln His Asp
10263           865           870           875           880
10264 Phe Gly Thr Ile Glu Ser Val Phe Glu Leu Ala Asn Arg Ser Lys Asp
10265           885           890           895
10266 Ile Asp Thr Leu Tyr Ala Asn Ser Gly Ala Gln Gly Arg Asp Leu Leu
10267           900           905           910
10268 Gln Thr Leu Leu Ile Asp Ser His Asp Ala Gly Tyr Ala Arg Gln Met
10269           915           920           925
10270 Ile Asp Asn Thr Ser Thr Gly Glu Ile Thr Lys Gln Leu Asn Ala Ala
10271           930           935           940
10272 Thr Thr Thr Leu Asn Asn Ile Ala Ser Leu Glu His Lys Thr Ser Ser
10273           945           950           955           960
10274 Leu Gln Thr Leu Ser Leu Ser Asn Ala Met Ile Leu Asn Ser Arg Leu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/881,752

DATE: 07/05/2001

TIME: 15:54:40

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

10275          965          970          975
10276 Val Asn Leu Ser Arg Arg His Thr Asn Asn Ile Asp Ser Phe Ala Gln
10277          980          985          990
10278 Arg Leu Gln Ala Leu Lys Asp Gln Lys Phe Ala Ser Leu Glu Ser Ala
10279          995          1000          1005
10280 Ala Glu Val Leu Tyr Gln Phe Ala Pro Lys Tyr Glu Lys Pro Thr Asn
10281          1010          1015          1020
10282 Val Trp Ala Asn Ala Ile Gly Gly Thr Ser Leu Asn Asn Gly Gly Asn
E--> 10283 (025) 1025          1030          1035          1040
10284 Ala Ser Leu Tyr Gly Thr Ser Ala Gly Val Asp Ala Tyr Leu Asn Gly
10285          1045          1050          1055
10286 Glu Val Glu Ala Ile Val Gly Gly Phe Gly Ser Tyr Gly Tyr Ser Ser
10287          1060          1065          1070
10288 Phe Asn Asn Gln Ala Asn Ser Leu Asn Ser Gly Ala Asn Asn Thr Asn
10289          1075          1080          1085
10290 Phe Gly Val Tyr Ser Arg Ile Phe Ala Asn Gln His Glu Phe Asp Phe
10291          1090          1095          1100
10292 Gly Ala Gln Gly Ala Leu Gly Ser Asp Gln Ser Ser Leu Asn Phe Lys
E--> 10293 (105)          1110          1115          1120
10294 Ser Ala Leu Leu Arg Asp Leu Asn Gln Ser Tyr Asn Tyr Leu Ala Tyr
10295          1125          1130          1135
10296 Ser Ala Ala Thr Arg Ala Ser Tyr Gly Tyr Asp Phe Ala Phe Phe Arg
10297          1140          1145          1150
10298 Asn Ala Leu Val Leu Lys Pro Ser Val Gly Val Ser Tyr Asn His Leu
10299          1155          1160          1165
10300 Gly Ser Thr Asn Phe Lys Ser Asn Ser Asn Gln Val Ala Leu Lys Asn
10301          1170          1175          1180
10302 Gly Ser Ser Ser Gln His Leu Phe Asn Ala Ser Ala Asn Val Glu Ala
E--> 10303 (185)          1190          1195          1200
10304 Arg Tyr Tyr Tyr Gly Asp Thr Ser Tyr Phe Tyr Met Asn Ala Gly Val
10305          1205          1210          1215
10306 Leu Gln Glu Phe Ala Asn Phe Gly Ser Ser Asn Ala Val Ser Leu Asn
10307          1220          1225          1230
10308 Thr Phe Lys Val Asn Ala Ala His Asn Pro Leu Ser Thr His Ala Arg
10309          1235          1240          1245
10310 Val Met Met Gly Gly Glu Leu Lys Leu Ala Lys Glu Val Phe Leu Asn
10311          1250          1255          1260
10312 Leu Gly Phe Val Tyr Leu His Asn Leu Ile Ser Asn Ile Gly His Phe
E--> 10313 (265)          1270          1275          1280
10314 Ala Ser Asn Leu Gly Met Arg Tyr Ser Phe
10315          1285          1290
11673 (2) INFORMATION FOR SEQ ID NO: 150:
11675 (i) SEQUENCE CHARACTERISTICS:
11676 (A) LENGTH: 1230 amino acids
11677 (B) TYPE: amino acid
11678 (C) STRANDEDNESS: single
11679 (D) TOPOLOGY: linear
11681 (ii) MOLECULE TYPE: protein
11682 (v) FRAGMENT TYPE: internal

```

When  
numbering  
the first  
amino acid  
on a line,  
begin the  
number under  
the first  
letter of the  
amino acid.  
Please leave  
a space  
after the  
ending number  
and the  
next amino  
acid.

11-12

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:41

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

11684      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:
11686 Met Ile Lys Lys Ala Arg Lys Phe Ile Pro Phe Phe Leu Ile Gly Ser
11687   1           5           10           15
11688 Leu Leu Ala Glu Asp Asn Gly Trp Tyr Met Ser Val Gly Tyr Gln Ile
11689           20           25           30
11690 Gly Gly Thr Gln Gln Phe Ile Asn Asn Lys Gln Leu Leu Glu Asn Gln
11691           35           40           45
11692 Asn Ile Ile Asn Ser Val Thr Gln Ser Ala Ile Asn Ile Ala Gly Pro
11693           50           55           60
11694 Thr Thr Gly Leu Ile Thr Leu Ser Ser Gln Thr Val Ile Asp Ala Leu
11695           65           70           75           80
11696 Gly Tyr Gly Val Ser Asn Thr Val Gly Asn Gln Leu Glu Gly Ile Ser
11697           85           90           95
11698 Asn Ile Leu Asn Gln Ile Gly Lys Arg Lys Asp Phe Tyr Ser Ser Arg
11699           100          105          110
11700 Gln Ile Ser Ser Ile Ser Gln Gln Ile Ile Gly Leu Lys Gly Ser Ser
11701           115          120          125
11702 Asp Pro Leu Lys Ala His Ser Ser Gln Ile Thr Ala Lys Leu Leu Ser
11703           130          135          140
11704 Asn Thr Gln Ser Ala Phe Asp Gln Gly Ile Ala Leu Ser Thr Asn Ile
11705           145          150          155          160
11706 Ile Ser Ser Ile Asn Ser Leu Asn Pro Ser Asn Asn Thr Gln Glu Val
11707           165          170          175
11708 Lys Lys Gln Leu Gln Asn Thr Ala Gln Ser Met Thr Glu Leu Leu Gln
11709           180          185          190
11710 Gln Ile Glu His Ser Ile Thr Lys Thr Thr Ser Thr Thr Tyr Ala Gln
11711           195          200          205
11712 Ser Leu Leu Ser Asn Leu Thr Asp Ala Val Asn Ala Ser Ser Asn Asn
11713           210          215          220
11714 Thr Ala Tyr Val Ser Ala Leu Val Asn Ala Leu Asn Thr Leu Gly Val
11715           225          230          235          240
11716 Gly Val Phe Pro Thr Thr Thr Thr Thr His Val Val Leu Asn Pro Pro
11717           245          250          255
11718 Gly Gln Val Val Phe Tyr Pro Thr Asn Ser Ile Leu Gly Ser Thr Ser
11719           260          265          270
11720 Ser Asn Ser Asn Asn Gln Gln Gln Tyr Asn Asn Thr Leu Leu Met Asn
11721           275          280          285
11722 Thr Leu Gln Gly Thr Leu Ser Ala Asn Thr Gln Asn Asn Pro Asn Gly
11723           290          295          300
11724 Cys Ala Asn Gln Val Gln Cys Leu Glu Gln Phe Ile Gln Asn Leu Ala
11725           305          310          315          320
11726 Pro Leu Ala Ala Thr Pro Thr Ser Asn Asn Gln Ala Asn Gln Gln Val
11727           325          330          335
11728 Gln Ala Ile Ala Gln Lys Leu Gln Ser Val Ala Ile Asn Thr Leu Asp
11729           340          345          350
11730 Asn Asn Ala Ile Asn Asn Thr Thr Tyr Asn Leu Asn Asn Leu His Asn
11731           355          360          365
11732 Ala Leu Asn Phe Gln Ala Tyr Glu Ser Thr Ile Glu Gln Tyr Asn Asn
11733           370          375          380

```

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:41

Input Set: A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

11734 Ala Leu Lys Gln Ile Ser Trp Ile Ser Phe Thr Glu Pro Lys Asn Leu
11735 385 390 395 400
11736 Leu Lys Asn Thr Ser Asn Asn Tyr Gln Ile Gly Thr Val Thr Asn Ala
11737 405 410 415
11738 Gln Gly Gln Asn Ile Ser Ala Tyr Asp Cys Met Thr Ala Thr Gly Ser
11739 420 425 430
11740 Leu Ser Ser Asn Ala Ser Ser Gly Ile Ser Cys Ser Ala Thr Ser Ser
11741 435 440 445
11742 Thr Ser Ser Thr Asn Ser Phe Asp Asn Ser Leu Val Ala Thr Ser Lys
11743 450 455 460
11744 Val Gln Thr Ile Asn Gly Lys Glu Gln Ile Gly Val Asn Ser Phe Asn
11745 465 470 475 480
11746 Leu Val Ser Gln Val Trp Ser Val Tyr Asn Ser Leu Lys Thr Ser Glu
11747 485 490 495
11748 Glu Asn Leu Gln Lys Asn Ala Asn Ile Leu Cys Ala Asn Gly Thr Gln
11749 500 505 510
11750 Ser Gly Thr Ser Ser Cys Asn Ser Ser Ser Gly Gly Leu Ser Ile Ser
11751 515 520 525
11752 Gly Asn Ala Gln Leu Gln Asn Ile Leu Ser Pro Thr Ser Gly Thr Thr
11753 530 535 540
11754 Thr Asn Thr Gln Ala Lys Ser Asn Ala Pro Lys Leu Lys Ala Met Val
11755 545 550 555 560
11756 Val Val Asn Asn Glu Glu Glu Ala Lys Thr Ala Asn Leu Ala Gln Ser
11757 565 570 575
11758 Ser Gly Thr Thr Thr Gln Ser Pro Asn Ser Thr Val Met Gly Ala Leu
11759 580 585 590
11760 Asn Thr Val Leu Gln Asn Val Ser Asn Phe Gln Gln Ser Ile Gln Asn
11761 595 600 605
11762 Ala Phe Gln Asn Gln Glu Ser Asn Ile Gln Ala Trp Ala Asn Ala Ile
11763 610 615 620
11764 Tyr Asn Thr Asn Gly Ser Gln Ser Gln Glu Met Thr Pro Asn Asn Asn
11765 625 630 635 640
11766 Gln Asp Leu Arg Ile Gln Leu Arg Ala Asn Phe Tyr Gln Leu Ile Asn
11767 645 650 655
11768 Thr Ile Asn Gln Gln Val Pro Thr Asp Met Asn Ala Leu Ile Asn Gln
11769 660 665 670
11770 Ser Gln Gln Thr Gln Gln Thr Ser Gly Ser Ala Ser Asn Asn Ala
11771 675 680 685
11772 Cys Ala Ser Gly Met Ser Gly Ser Asn Gly Asn Trp Cys Tyr Gln Gln
11773 690 695 700
11774 Trp Ser Asp Ser Lys Ala Tyr Tyr Ser Gly Leu Gln Ser Ala Leu Gly
11775 705 710 715 720
11776 Tyr Gln Thr Gln Ala Thr Thr Gln Ser Gly Ser Asn Gly Gly Asn Ser
11777 725 730 735
11778 Ile Thr Tyr Asn Val Gln Gln Ile Thr Leu Thr Ser Asn Gly Leu Leu
11779 740 745 750
11780 Asn Gln Ile Ile Thr Asn Leu Lys Ser Val Asn Gly Gly Asn Gly Ala
11781 755 760 765
11782 Ser Gly Thr Gly Ser Gly Asn Gly Thr Ser Gln Ile Asn Thr Ala Tyr

```

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:41

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

11783      770      775      780
11784 Gln Met Leu Thr Asp Ala Ser Asp Gly Lys Leu Gly Thr Tyr Ser Ser
11785 785      790      795      800
11786 Ser Ser Gly Ser Asn Asn Gly Tyr Thr Pro Cys Asn Ser Thr Asn Gly
11787      805      810      815
11788 Ser Asn Lys Thr Ser Gly Asn Asn Cys Tyr Glu Pro Asn Lys Gln Gln
11789      820      825      830
11790 Asn Ala Thr Thr Ala Thr Ala Thr Thr Asp Ser Asn Leu Gln Lys Val
11791      835      840      845
11792 Tyr Asn Asp Ala Gln Lys Ile Ala Asn Ile Ile Ala Ser Ser Gly Asn
11793      850      855      860
11794 Asn Lys Gly Val Glu Asn Gly Leu Lys Gln Phe Phe Glu Ala Leu Lys
11795 865      870      875      880
11796 Asn Asn Ser Ser Ser Leu Ser Asn Leu Cys Gly Asn Gly Ser Ser Gly
11797      885      890      895
11798 Ser Ser Gly Thr Thr Cys Ser Gly Trp Leu Ile Asn Leu Leu Gly Ala
11799      900      905      910
11800 Ile Pro Thr Asn Gly Val Ser Asp Thr Asn Asn Leu Ile Asn Leu Leu
11801      915      920      925
11802 Thr Glu Phe Ile Lys Thr Ala Gly Phe Ile Gln Asn Asn Asp Ser Ser
11803      930      935      940
11804 Val Ser Thr Ser Leu Thr Ser Ala Phe Gln Ala Ile Thr Ser Ala Ile
11805 945      950      955      960
11806 Ser Gln Gly Phe Gln Ala Leu Gln Asn Asp Ile Ser Pro Asn Ala Ile
11807      965      970      975
11808 Leu Thr Leu Leu Gln Glu Ile Thr Ser Asn Thr Thr Thr Ile Gln Ser
11809      980      985      990
11810 Phe Ser Gln Thr Leu Arg Gln Leu Leu Gly Asp Lys Thr Phe Phe Met
11811      995      1000      1005
11812 Ala Gln Gln Lys Leu Ile Asp Ala Met Ile Asn Ala Arg Asn Gln Val
11813      1010      1015      1020
11814 Gln Asn Ala Gln Asn Gln Ala Asn Asn Tyr Gly Ser Gln Pro Val Leu
E--> 11815 025      1030      1035      1040
11816 Ser Gln Tyr Ala Ala Lys Ser Thr Gln His Gly Met Ser Asn Gly
11817      1045      1050      1055
11818 Leu Gly Val Gly Leu Gly Tyr Lys Tyr Phe Phe Gly Lys Ala Arg Lys
11819      1060      1065      1070
11820 Leu Gly Leu Arg His Tyr Phe Phe Phe Asp Tyr Gly Phe Ser Glu Ile
11821      1075      1080      1085
11822 Gly Leu Ala Asn Gln Ser Val Lys Ala Asn Ile Phe Ala Tyr Gly Val
11823      1090      1095      1100
11824 Gly Thr Asp Phe Leu Trp Asn Leu Phe Arg Arg Thr Tyr Asn Thr Lys
E--> 11825 105      1110      1115      1120
11826 Ala Leu Asn Phe Gly Leu Phe Ala Gly Val Gln Leu Gly Gly Ala Thr
11827      1125      1130      1135
11828 Trp Leu Ser Ser Leu Arg Gln Gln Ile Ile Asp Asn Trp Gly Ser Ala
11829      1140      1145      1150
11830 Asn Asp Ile His Ser Thr Asn Phe Gln Val Ala Leu Asn Phe Gly Val
11831      1155      1160      1165

```

*same  
error*

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/881,752

DATE: 07/05/2001

TIME: 15:54:41

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

11832 Arg Thr Asn Phe Ala Glu Phe Lys Arg Phe Ala Lys Lys Phe His Asn  
 11833 1170 1175 1180  
 11834 Gln Gly Val Ile Ser Gln Lys Ser Val Glu Phe Gly Ile Lys Val Pro  
 E--> 11835 185 1190 1195 1200 *same*  
 11836 Leu Ile Asn Gln Ala Tyr Leu Asn Ser Ala Gly Ala Asp Val Ser Tyr  
 11837 1205 1210 1215  
 11838 Arg Arg Leu Tyr Thr Phe Tyr Ile Asn Tyr Ile Met Gly Phe  
 11839 1220 1225 1230  
 23321 (2) INFORMATION FOR SEQ ID NO: 308:  
 23323 (i) SEQUENCE CHARACTERISTICS:  
 23324 (A) LENGTH: 1797 amino acids  
 23325 (B) TYPE: amino acid  
 23326 (C) STRANDEDNESS: single  
 23327 (D) TOPOLOGY: linear  
 23329 (ii) MOLECULE TYPE: protein  
 23330 (v) FRAGMENT TYPE: internal  
 23332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:  
 23334 Met Ala Gly Arg His Gly Asn Lys Gly Ile Val Ser Asn Ile Val Pro  
 23335 1 5 10 15  
 23336 Val Ala Asp Met Pro Tyr Thr Ala Asp Gly Glu Pro Val Asp Ile Val  
 23337 20 25 30  
 23338 Leu Asn Pro Leu Gly Val Pro Ser Arg Met Asn Ile Gly Gln Ile Leu  
 23339 35 40 45  
 23340 Glu Met His Leu Gly Leu Val Gly Lys Glu Phe Gly Lys Gln Ile Ala  
 23341 50 55 60  
 23342 Arg Met Leu Glu Asp Lys Thr Lys Asp Phe Ala Lys Glu Leu Arg Ala  
 23343 65 70 75 80  
 23344 Lys Met Leu Glu Xaa Ala Asn Ala Ile Asn Glu Lys Asp Pro Leu Thr  
 23345 85 90 95  
 23346 Ile His Ala Leu Glu Asn Cys Ser Asp Glu Glu Leu Leu Glu Tyr Ala  
 23347 100 105 110  
 23348 Lys Asp Trp Ser Lys Gly Val Lys Met Ala Ile Pro Val Phe Glu Gly  
 23349 115 120 125  
 23350 Ile Ser Gln Glu Lys Phe Tyr Lys Leu Phe Glu Leu Ala Lys Ile Ala  
 23351 130 135 140  
 23352 Met Asp Gly Lys Met Asp Leu Tyr Asp Gly Arg Thr Gly Glu Lys Met  
 23353 145 150 155 160  
 23354 Arg Glu Arg Val Asn Val Gly Tyr Met Tyr Met Ile Lys Leu His His  
 23355 165 170 175  
 23356 Leu Val Asp Glu Lys Val His Ala Arg Ser Thr Gly Pro Tyr Ser Leu  
 23357 180 185 190  
 23358 Val Thr His Gln Pro Val Gly Gly Lys Ala Leu Phe Gly Gly Gln Arg  
 23359 195 200 205  
 23360 Phe Gly Glu Met Glu Val Trp Ala Leu Glu Ala Tyr Gly Ala Ala His  
 23361 210 215 220  
 23362 Thr Leu Lys Glu Met Leu Thr Ile Lys Ser Asp Asp Ile Arg Gly Arg  
 23363 225 230 235 240  
 23364 Glu Asn Ala Tyr Arg Ala Ile Ala Lys Gly Glu Gln Val Gly Glu Ser  
 23365 245 250 255

pr 14-16  
 WLS

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:43

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

23366  Glu Ile Pro Glu Thr Phe Tyr Val Leu Thr Lys Glu Leu Gln Ser Leu
23367                260                      265                      270
23368  Ala Leu Asp Ile Asn Ile Phe Gly Asp Asp Val Asp Glu Asp Gly Ala
23369                275                      280                      285
23370  Pro Lys Pro Ile Val Ile Lys Glu Asp Asp Arg Pro Lys Asp Phe Ser
23371                290                      295                      300
23372  Ser Phe Gln Leu Thr Leu Ala Ser Pro Glu Lys Ile His Ser Trp Ser
23373                305                      310                      315                      320
23374  Tyr Gly Glu Val Lys Lys Pro Glu Thr Ile Asn Tyr Arg Thr Leu Lys
23375                325                      330                      335
23376  Pro Glu Arg Asp Gly Leu Phe Cys Met Lys Ile Phe Gly Pro Thr Lys
23377                340                      345                      350
23378  Asp Tyr Glu Cys Leu Cys Gly Lys Tyr Lys Lys Pro Arg Phe Lys Asp
23379                355                      360                      365
23380  Ile Gly Thr Cys Glu Lys Cys Gly Val Ala Ile Thr His Ser Lys Val
23381                370                      375                      380
23382  Arg Arg Phe Arg Met Gly His Ile Glu Leu Ala Thr Pro Val Ala His
23383                385                      390                      395                      400
23384  Ile Trp Tyr Val Asn Ser Leu Pro Ser Arg Ile Gly Thr Leu Leu Gly
23385                405                      410                      415
23386  Val Lys Met Lys Asp Leu Glu Arg Val Leu Tyr Tyr Glu Ala Tyr Ile
23387                420                      425                      430
23388  Val Lys Glu Pro Gly Glu Ala Ala Tyr Asp Asn Glu Gly Thr Lys Leu
23389                435                      440                      445
23390  Val Met Lys Tyr Asp Ile Leu Asn Glu Glu Gln Tyr Gln Asn Ile Ser
23391                450                      455                      460
23392  Arg Arg Tyr Glu Asp Arg Gly Phe Val Ala Gln Met Gly Gly Glu Ala
23393                465                      470                      475                      480
23394  Ile Lys Asp Leu Leu Glu Glu Ile Asp Leu Ile Thr Leu Leu Gln Ser
23395                485                      490                      495
23396  Leu Lys Glu Glu Val Lys Asp Thr Asn Ser Asp Ala Lys Lys Lys Lys
23397                500                      505                      510
23398  Leu Ile Lys Arg Leu Lys Val Val Glu Ser Phe Leu Asn Ser Gly Asn
23399                515                      520                      525
23400  Arg Pro Glu Trp Met Met Leu Thr Val Leu Pro Val Leu Pro Pro Asp
23401                530                      535                      540
23402  Leu Arg Pro Leu Val Ala Leu Asp Gly Gly Lys Phe Ala Val Ser Asp
23403                545                      550                      555                      560
23404  Val Asn Glu Leu Tyr Arg Arg Val Ile Asn Arg Asn Gln Arg Leu Lys
23405                565                      570                      575
23406  Arg Leu Met Glu Leu Gly Ala Pro Glu Ile Ile Val Arg Asn Glu Lys
23407                580                      585                      590
23408  Arg Met Leu Gln Glu Ala Val Asp Val Leu Phe Asp Asn Gly Arg Ser
23409                595                      600                      605
23410  Thr Asn Ala Val Lys Gly Ala Asn Lys Arg Pro Leu Lys Ser Leu Ser
23411                610                      615                      620
23412  Glu Ile Ile Lys Gly Lys Gln Gly Arg Phe Arg Gln Asn Leu Leu Gly
23413                625                      630                      635                      640
23414  Lys Arg Val Asp Phe Ser Gly Arg Ser Val Ile Val Val Gly Pro Asn

```

## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:43

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

23415					645					650					655	
23416	Leu	Lys	Met	Asp	Glu	Cys	Gly	Leu	Pro	Lys	Asn	Met	Ala	Leu	Glu	Leu
23417					660					665					670	
23418	Phe	Lys	Pro	His	Leu	Leu	Ser	Lys	Leu	Glu	Glu	Arg	Gly	Tyr	Ala	Thr
23419					675					680					685	
23420	Thr	Leu	Lys	Gln	Ala	Lys	Arg	Met	Ile	Glu	Gln	Lys	Ser	Asn	Glu	Val
23421					690					695					700	
23422	Trp	Glu	Cys	Leu	Gln	Glu	Ile	Thr	Glu	Gly	Tyr	Pro	Val	Leu	Leu	Asn
23423					705					710					715	
23424	Arg	Ala	Pro	Thr	Leu	His	Lys	Gln	Ser	Ile	Gln	Ala	Phe	His	Pro	Lys
23425					725					730					735	
23426	Leu	Ile	Asp	Gly	Lys	Ala	Ile	Gln	Leu	His	Pro	Leu	Val	Cys	Ser	Ala
23427					740					745					750	
23428	Phe	Asn	Ala	Asp	Phe	Asp	Gly	Asp	Gln	Met	Ala	Val	His	Val	Pro	Leu
23429					755					760					765	
23430	Ser	Gln	Glu	Ala	Ile	Ala	Glu	Cys	Lys	Val	Leu	Met	Leu	Ser	Ser	Met
23431					770					775					780	
23432	Asn	Ile	Leu	Leu	Pro	Ala	Ser	Gly	Lys	Ala	Val	Ala	Ile	Pro	Ser	Gln
23433					785					790					795	
23434	Asp	Met	Val	Leu	Gly	Leu	Tyr	Tyr	Leu	Ser	Leu	Glu	Lys	Ser	Gly	Val
23435					805					810					815	
23436	Lys	Gly	Glu	His	Lys	Leu	Phe	Ser	Ser	Val	Asn	Glu	Ile	Ile	Thr	Ala
23437					820					825					830	
23438	Ile	Asp	Thr	Lys	Glu	Leu	Asp	Ile	His	Ala	Lys	Ile	Arg	Val	Leu	Asp
23439					835					840					845	
23440	Gln	Gly	Asn	Ile	Ile	Ala	Thr	Ser	Ala	Gly	Arg	Met	Ile	Ile	Lys	Ser
23441					850					855					860	
23442	Ile	Leu	Pro	Asp	Phe	Ile	Pro	Thr	Asp	Leu	Trp	Asn	Arg	Pro	Met	Lys
23443					865					870					875	
23444	Lys	Lys	Asp	Ile	Gly	Val	Leu	Val	Asp	Tyr	Val	His	Lys	Val	Gly	Gly
23445					885					890					895	
23446	Ile	Gly	Ile	Thr	Ala	Thr	Phe	Leu	Asp	Asn	Leu	Lys	Thr	Leu	Gly	Phe
23447					900					905					910	
23448	Arg	Tyr	Ala	Thr	Lys	Ala	Gly	Ile	Ser	Ile	Ser	Met	Glu	Asp	Ile	Ile
23449					915					920					925	
23450	Thr	Pro	Lys	Asp	Lys	Gln	Lys	Met	Val	Glu	Lys	Ala	Lys	Val	Glu	Val
23451					930					935					940	
23452	Lys	Lys	Ile	Gln	Gln	Gln	Tyr	Asp	Gln	Gly	Leu	Leu	Thr	Asp	Gln	Glu
23453					945					950					955	
23454	Arg	Tyr	Asn	Lys	Ile	Ile	Asp	Thr	Trp	Thr	Glu	Val	Asn	Asp	Lys	Met
23455					965					970					975	
23456	Ser	Lys	Glu	Met	Met	Thr	Ala	Ile	Ala	Gln	Asp	Lys	Glu	Gly	Phe	Asn
23457					980					985					990	
23458	Ser	Ile	Tyr	Met	Met	Ala	Asp	Ser	Gly	Ala	Arg	Gly	Ser	Ala	Ala	Gln
23459					995					1000					1005	
23460	Ile	Arg	Gln	Leu	Ser	Ala	Met	Arg	Gly	Leu	Met	Thr	Lys	Pro	Asp	Gly
23461					1010					1015					1020	
23462	Ser	Ile	Ile	Glu	Thr	Pro	Ile	Ile	Ser	Asn	Phe	Lys	Glu	Gly	Leu	Asn
E--> 23463					1025					1030					1035	
															1040	

*Same*



## RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:43

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

23464 Val Leu Glu Tyr Phe Asn Ser Thr His Gly Ala Arg Lys Gly Leu Ala  
 23465 1045 1050 1055  
 23466 Asp Thr Ala Leu Lys Thr Ala Asn Ala Gly Tyr Leu Thr Arg Lys Leu  
 23467 1060 1065 1070  
 23468 Ile Asp Val Ser Gln Asn Val Lys Val Val Ser Asp Asp Cys Gly Thr  
 23469 1075 1080 1085  
 23470 His Glu Gly Ile Glu Ile Thr Asp Ile Ala Val Gly Ser Glu Leu Ile  
 23471 1090 1095 1100  
 23472 Glu Pro Leu Glu Glu Arg Ile Phe Gly Arg Val Leu Leu Glu Asp Val  
 E--> 23473 105 1110 1115 1120  
 23474 Ile Asp Pro Ile Thr Asn Glu Ile Leu Leu Tyr Ala Asp Thr Leu Ile  
 23475 1125 1130 1135  
 23476 Asp Glu Glu Gly Ala Lys Lys Val Val Glu Ala Gly Ile Lys Ser Ile  
 23477 1140 1145 1150  
 23478 Thr Ile Arg Thr Pro Val Thr Cys Lys Ala Pro Lys Gly Val Cys Ala  
 23479 1155 1160 1165  
 23480 Lys Cys Tyr Gly Leu Asn Leu Gly Glu Gly Lys Met Ser Tyr Pro Gly  
 23481 1170 1175 1180  
 23482 Glu Ala Val Gly Val Val Ala Ala Gln Ser Ile Gly Glu Pro Gly Thr  
 E--> 23483 185 1190 1195 1200  
 23484 Gln Leu Thr Leu Arg Thr Phe His Val Gly Gly Thr Ala Ser Arg Ser  
 23485 1205 1210 1215  
 23486 Gln Asp Glu Arg Glu Ile Val Ala Ser Lys Glu Gly Phe Val Arg Phe  
 23487 1220 1225 1230  
 23488 Tyr Asn Leu Arg Thr Tyr Thr Asn Lys Glu Gly Lys Asn Ile Ile Ala  
 23489 1235 1240 1245  
 23490 Asn Arg Arg Asn Ala Ser Ile Leu Val Val Glu Pro Lys Ile Lys Ala  
 23491 1250 1255 1260  
 23492 Pro Phe Asp Gly Glu Leu Arg Ile Glu Thr Val Tyr Glu Glu Val Val  
 E--> 23493 265 1270 1275 1280  
 23494 Val Ser Val Lys Asn Gly Asp Gln Glu Ala Lys Phe Val Leu Arg Arg  
 23495 1285 1290 1295  
 23496 Ser Asp Ile Val Lys Pro Ser Glu Leu Ala Gly Val Gly Gly Lys Ile  
 23497 1300 1305 1310  
 23498 Glu Gly Lys Val Tyr Leu Pro Tyr Ala Ser Gly His Lys Val His Lys  
 23499 1315 1320 1325  
 23500 Gly Gly Ser Ile Ala Asp Ile Ile Gln Glu Gly Trp Asn Val Pro Asn  
 23501 1330 1335 1340  
 23502 Arg Ile Pro Tyr Ala Ser Glu Leu Leu Val Lys Asp Asn Asp Pro Ile  
 E--> 23503 345 1350 1355 1360  
 23504 Ala Gln Asp Val Tyr Ala Lys Glu Lys Gly Val Ile Lys Tyr Tyr Val  
 23505 1365 1370 1375  
 23506 Leu Glu Ala Asn His Leu Glu Arg Thr His Gly Ile Lys Lys Gly Asp  
 23507 1380 1385 1390  
 23508 Met Val Ser Glu Lys Gly Leu Phe Ala Val Ile Ala Asp Asp Asn Gly  
 23509 1395 1400 1405  
 23510 Arg Glu Ala Ala Arg His Tyr Ile Ala Arg Gly Ser Glu Ile Leu Ile  
 23511 1410 1415 1420  
 23512 Asp Asp Asn Ser Glu Val Ser Thr Asn Ser Val Ile Ser Lys Pro Thr

*Same*

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/881,752

DATE: 07/05/2001

TIME: 15:54:43

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

```

E--> 23513 425          1430          1435          1440
23514 Thr Asn Thr Phe Lys Thr Ile Ala Thr Trp Asp Pro Tyr Asn Thr Pro
23515          1445          1450          1455
23516 Ile Ile Ala Asp Phe Lys Gly Lys Val Gly Phe Val Asp Val Ile Ala
23517          1460          1465          1470
23518 Gly Val Thr Val Ala Glu Lys Glu Asp Glu Asn Thr Gly Ile Thr Ser
23519          1475          1480          1485
23520 Leu Val Val Asn Asp Tyr Ile Pro Ser Gly Tyr Lys Pro Ser Leu Phe
23521          1490          1495          1500
23522 Leu Glu Gly Ala Asn Gly Glu Glu Met Arg Tyr Phe Leu Glu Pro Lys
E--> 23523 505          1510          1515          1520
23524 Thr Ser Ile Ala Ile Ser Asp Gly Ser Ser Val Glu Gln Ala Glu Val
23525          1525          1530          1535
23526 Leu Ala Lys Ile Pro Lys Ala Thr Val Lys Ser Arg Asp Ile Thr Gly
23527          1540          1545          1550
23528 Gly Leu Pro Arg Val Ser Glu Leu Phe Glu Ala Arg Lys Pro Lys Pro
23529          1555          1560          1565
23530 Lys Asp Val Ala Ile Leu Ser Glu Val Asp Gly Ile Val Ser Phe Gly
23531          1570          1575          1580
23532 Lys Pro Ile Arg Asn Lys Glu His Ile Ile Val Thr Ser Lys Asp Gly
E--> 23533 585          1590          1595          1600
23534 Arg Ser Met Asp Tyr Phe Val Asp Lys Gly Lys Gln Ile Leu Val His
23535          1605          1610          1615
23536 Ala Asp Glu Phe Val His Ala Gly Glu Ala Met Thr Asp Gly Val Ile
23537          1620          1625          1630
23538 Ser Ser His Asp Ile Leu Arg Ile Ser Gly Glu Lys Glu Leu Tyr Lys
23539          1635          1640          1645
23540 Tyr Ile Val Ser Glu Val Gln Gln Val Tyr Arg Arg Gln Gly Val Ser
23541          1650          1655          1660
23542 Ile Ala Asp Lys His Ile Glu Ile Ile Val Ser Gln Met Leu Arg Gln
E--> 23543 665          1670          1675          1680
23544 Val Arg Ile Leu Asp Ser Gly Asp Ser Lys Phe Ile Glu Gly Asp Leu
23545          1685          1690          1695
23546 Val Ser Lys Lys Leu Phe Lys Glu Glu Asn Ala Arg Val Ile Ala Leu
23547          1700          1705          1710
23548 Lys Gly Glu Pro Ala Ile Ala Glu Pro Val Leu Leu Gly Ile Thr Arg
23549          1715          1720          1725
23550 Ala Ala Ile Gly Ser Asp Ser Ile Ile Ser Ala Ala Ser Phe Gln Glu
23551          1730          1735          1740
23552 Thr Thr Lys Val Leu Thr Glu Ala Ser Ile Ala Met Lys Lys Asp Phe
E--> 23553 745          1750          1755          1760
23554 Leu Glu Asp Leu Lys Glu Asn Val Val Leu Gly Arg Met Ile Pro Val
23555          1765          1770          1775
23556 Gly Thr Gly Met Tyr Lys Asn Lys Lys Ile Val Leu Arg Ala Leu Glu
23557          1780          1785          1790
23558 Asp Asn Ser Lys Phe
23559          1795

```

## VERIFICATION SUMMARY

DATE: 07/05/2001

PATENT APPLICATION: US/09/881,752

TIME: 15:54:45

Input Set : A:\06132.041001 sequence listing.txt

Output Set: N:\CRF3\07032001\I881752.raw

L:3 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]  
L:7 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]  
L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:1332 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19  
L:1358 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20  
L:1522 M:254 E: No. of Bases conflict, Input:575 Counted:1815 SEQ:21  
L:1522 M:320 E: (1) Wrong Nucleic Acid Designator, 20  
L:1522 M:111 C: (47) String data converted to upper case,  
M:254 Repeated in SeqNo=21  
L:1526 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21  
L:1530 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21  
L:1534 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21  
L:1538 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21  
L:1542 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21  
L:1546 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21  
L:1549 M:204 E: No. of Bases differ, LENGTH:Input:2169 Counted:2217 SEQ:21  
L:2276 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29  
L:2308 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30  
L:5730 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77  
L:5780 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78  
L:6526 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85  
L:6585 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86  
L:8330 M:333 E: Wrong sequence grouping, Nucleotides not in groups!  
L:9001 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:125  
L:9027 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:126  
L:9078 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:127  
L:9206 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:128  
L:10283 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:138  
M:332 Repeated in SeqNo=138  
L:11815 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:150  
M:332 Repeated in SeqNo=150  
L:15906 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:203  
L:16012 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:204  
L:19424 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255  
L:19498 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:256  
L:22418 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:299  
L:22492 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:300  
L:22888 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:307  
L:23344 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:308  
L:23463 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:308  
M:332 Repeated in SeqNo=308